

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2003, 22:00:37 ; Search time 37 Seconds
(without alignments)
1884.108 Million cell updates/sec

Title: US-09-815-923-4

Perfect score: 3141
Sequence: 1 MPSPDAPPATAPPPDLPAT.....TIQREPVTSIPPADSTLCNL 587

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Minimum number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA: *
1: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep: *
2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep: *
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep: *
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep: *
5: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep: *
6: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep: *
7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep: *
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep: *
9: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep: *
10: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep: *
11: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep: *
12: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep: *
13: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep: *
14: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep: *
15: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep: *
16: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep: *
17: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3141	100.0	587	11	US-09-815-923-4
2	2333.5	74.3	622	10	US-09-843-598-11
3	1726	55.0	630	10	US-09-843-598-10
4	1579.5	50.3	671	10	US-09-843-598-5
5	1571.5	50.0	671	10	US-09-843-598-7
6	1532.5	48.8	624	10	US-09-795-693-24
7	1532.5	48.8	624	15	US-10-156-239-24
8	1532.5	48.8	624	15	US-10-156-239-24
9	1369.5	43.6	797	10	US-09-795-485-24
10	1302	41.5	597	11	US-09-815-923-2
11	1247.5	39.7	614	12	US-09-919-039-378
12	1221	38.9	727	11	US-09-815-923-10
13	1077	34.3	421	10	US-09-843-598-9
14	1069	34.0	556	11	US-09-815-923-6
15	990	31.5	437	11	US-09-818-656A-4
16	990	31.5	437	15	US-10-216-441-4

17	990	31.5	459	11	US-09-818-656A-2	Sequence 2, Appli
18	990	31.5	459	15	US-10-216-441-2	Sequence 2, Appli
19	981	31.2	730	10	US-09-741-149-2	Sequence 2, Appli
20	981	31.2	730	10	US-09-795-693-5	Sequence 5, Appli
21	981	31.2	730	15	US-10-156-239-5	Sequence 5, Appli
22	981	31.2	730	15	US-10-199-485-5	Sequence 5, Appli
23	973	31.0	729	10	US-09-741-149-4	Sequence 4, Appli
24	964.5	30.7	599	11	US-09-861-846-4	Sequence 4, Appli
25	960	30.6	676	11	US-09-815-923-12	Sequence 12, Appli
26	943.5	30.0	610	11	US-09-861-846-2	Sequence 2, Appli
27	926.5	29.5	727	10	US-09-923-444A-2	Sequence 2, Appli
28	618	19.7	579	11	US-09-738-626-4648	Sequence 4648, Ap
29	461.5	14.7	224	10	US-09-843-598-6	Sequence 6, Appli
30	300	9.6	449	12	US-09-769-787-110	Sequence 110, App
31	167	5.3	196	12	US-09-989-442-87	Sequence 87, Appli
32	142	4.5	52	10	US-09-864-761-43363	Sequence 43363, A
33	136	4.3	1098	11	US-09-712-363-288	Sequence 288, App
34	133.5	4.3	84	10	US-09-864-761-39755	Sequence 39755, A
35	118.5	3.8	77	10	US-09-864-761-39850	Sequence 39850, A
36	117	3.7	403	15	US-10-140-372-2	Sequence 2, Appli
37	116	3.7	489	15	US-10-156-761-8760	Sequence 8760, Ap
38	111.5	3.5	442	15	US-10-156-761-9153	Sequence 9153, Ap
39	111	3.5	475	10	US-09-815-242-11503	Sequence 11503, A
40	110	3.5	505	12	US-09-769-787-11388	Sequence 120, App
41	109.5	3.5	519	10	US-09-815-242-11388	Sequence 11388, A
42	109.5	3.5	519	11	US-09-895-913A-118	Sequence 118, App
43	109.5	3.5	1083	14	US-10-080-170-330	Sequence 330, App
44	109	3.5	553	12	US-09-836-705-38	Sequence 38, Appli
45	108	3.4	510	10	US-09-815-242-13607	Sequence 13607, A

ALIGNMENTS

RESULT 1
US-09-815-923-4
; Sequence 4, Application US/09815923
; Publication No. US20020197644A1
; GENERAL INFORMATION:
; APPLICANT: Sarjeet S.
; APPLICANT: Ross, Linda S.
; TITLE OF INVENTION: The Regents of the University of California
; TITLE OF INVENTION: Use of Insect Cell Membrane Transporters as No. US20020197644A
; FILE REFERENCE: 023070-093800US
; CURRENT APPLICATION NUMBER: US/09/815,923
; CURRENT FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 587
; TYPE: PRT
; ORGANISM: Manduca sexta
; FEATURE:
; OTHER INFORMATION: serotonin transporter
US-09-815-923-4

Query Match 100.0%: Score 3141; DB 11; Length 587;
Best Local Similarity 100.0%: Pred. No. 6.7e-274;
Matches 587; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPSPDAPPATAPPPDLPATTAOKSRVSVSLTPARQRETMWAKAEFLLAVFADLGN 60
1 VWFPPYICVONGGAEFLIPYCVMLFPGDLPLFLELALGQYHRGGLTLMKRCPALRGV 120
Db 1 MPSPDAPPATAPPPDLPATTAOKSRVSVSLTPARQRETMWAKAEFLLAVFADLGN 60
QY 61 VWFPPYICVONGGAEFLIPYCVMLFPGDLPLFLELALGQYHRGGLTLMKRCPALRGV 120
Db 61 VWFPPYICVONGGAEFLIPYCVMLFPGDLPLFLELALGQYHRGGLTLMKRCPALRGV 120
QY 121 GYAIICMIDIMGMVYNTIIGWAVYILIASIASINSVLPMTSCDNEWNTPLCTPVTSPQTN 180
Db 121 GYAIICMIDIMGMVYNTIIGWAVYILIASIASINSVLPMTSCDNEWNTPLCTPVTSPQTN 180

OY	181	NNSSPAKEPFRNVL	EOHKSNGLDMG	IKRSLLCYGVVLYEFS	LWKGRSAGKV	240
Db	181	PNSSTPAKEPFRNVL	EOHKSNGLDMG	IKRSLLCYGVVLYEFS	LWKGRSAGKV	240
OY	241	WVTLAPYVVL	LLILARGVTL	PGATEGIRYVYL	PEPHKLQNSK	WIDAA
Db	241	WVTLAPYVVL	LLILARGVTL	PGATEGIRYVYL	PEPHKLQNSK	WIDAA
OY	301	GTLLASSTYK	FNNNCYDALIT	SSINCLTSL	LAGVLFVS	UYMAHONK
Db	301	GTLLASSTYK	FNNNCYDALIT	SSINCLTSL	LAGVLFVS	UYMAHONK
OY	361	PGLVFIYVPE	AIATMGSVFMA	IFPFMLITGL	DSTFEGGLE	FAVYAL
Db	361	PGLVFIYVPE	AIATMGSVFMA	IFPFMLITGL	DSTFEGGLE	FAVYAL
OY	421	EVEFVAVLL	LFYICALPTTY	GGVLYVDL	LVNYGGL	ALLFVFAE
Db	421	EVEFVAVLL	LFYICALPTTY	GGVLYVDL	LVNYGGL	ALLFVFAE
OY	481	EDVKTMLCHT	PGMEFWKCM	SYISPVFL	LVLEFVS	LAHEML
Db	481	EDVKTMLCHT	PGMEFWKCM	SYISPVFL	LVLEFVS	LAHEML
OY	541	GTVYSCIP	LYITIKLLIT	PGNCINIK	TQREVN	SIPADSTLC
Db	541	GTVYSCIP	LYITIKLLIT	PGNCINIK	TQREVN	SIPADSTLC

RESULT 2
US-09-843-598-11
; Sequence 11, Application US/09843598
; Patent No. US20020010944A1

; GENERAL INFORMATION:

```

: APPLICANT: Horvitz, H. Robert
: APPLICANT: Ranganathan, Rajesh
: TITLE OF INVENTION: CESTR GENES, PROTEINS, AND MODULATORY
: TITLE OF INVENTION: COMPOUNDS
: FILE REFERENCE: 01597/525002
: CURRENT APPLICATION NUMBER: US/09/843,598
: CURRENT FILING DATE: 2001-04-26
: PRIOR APPLICATION NUMBER: US 60/200,549
: PRIOR FILING DATE: 2000-04-26
: NUMBER OF SEQ ID NOS: 11
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 11
: LENGTH: 622
: TYPE: prt
: ORGANISM: Drosophila melanogaster
: 09-843-598-11

```

Query Match	74.3%	Score	2333.5	DB	10	Length	622
Best Local Similarity	76.58%	Pred. No.	3e-201				
Matches	434	Conservative	50	Mismatches	76	Indels	7
						Gaps	4

OY	8	PAPAPRPDDLATATQKRSRVVSLTPRORRETKAKMKELLAAVGAADLGNWFEPTI	67
Dd	47	PAKYTD--LAPKLANNERILVYST-ERTRETQOKAEFLAVIGFAVDLGNWFEPIT	103
OY	68	CYONGGAFLLPYCYMLLPGLPLPFFLELAGOYHRCGLTLMKRICPALKGVAICMI	127
Dd	104	CYONGGAFVLPYCFLTFELFGLPLPYEMETALQOFHRCGLSTLMKRICPALKGVAICTI	163
OY	128	DIYMGMYNTTIYGNAVYYLIASINSVLPMPTSCDNENMTPLCPVTSPOTNPNSSIPA	187
Dd	164	DIYMGMYNTTIIGNAVYYLT--FAEFTSKLPTSCDNPINTEENCHQVTSENTPELATSPA	220
OY	188	KEFERANVLBOHKSGLDMPGRKISLACGVGAVLYEESLMKGYRSAGKVWTATLAP	247
Dd	221	KEFERAKVLESYKGGDLDFMGPVKVTTLACVCGAVLYEFSLMKGYRSAGKVWWTATLAP	280
OY	248	YVVLLILLIARGVTLPGATEGIIRYVLTPEMHKLONSKWVIDAASQIFESLGPGFGTLLAS	307

[illegible]

RESULT 3
US-09-843-598-10

; Sequence 10, Application US/09843598
; Patent No. US2002001094A1

; GENERAL INFORMATION:

```

:
: APPLICANT: Horvitz, H. Robert
: APPLICANT: Ranganathan, Rajesh
: TITLE OF INVENTION: CESTERT GENES, PROTEINS, AND MODULATORY
: TITLE OF INVENTION: COMPOUNDS
: FILE REFERENCE: 01997/525002
: CURRENT APPLICATION NUMBER: US/09/843,598
: CURRENT FILING DATE: 2001-04-26
: PRIOR APPLICATION NUMBER: US 60/200,549
: PRIOR FILING DATE: 2000-04-26
: NUMBER OF SEQ ID NOS: 11
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 10
: LENGTH: 630
: TYPE: PRN
: ORGANISM: Homo sapiens
US-09-843-598-10

```

Query Match	55.0%	Score 1726;	DB 10	Length 630;
Best Local Similarity	54.5%	Pred. NO. 1.3e-146;		
Matches 320; Conservative	97;	Mismatches 154;	Indels 16;	Gaps 6

```

0Y 6 PPAAPTAPP---DLPAATACKSRSVVSLTPARQREIMAKKAPELLAVGVADLVNW 62
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 49 AVPSGADDDTRHSIPATP-----TTVAELDHQGERETWGGKKVDLVSIGYADVLGNW 103
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
0Y 63 RPPYCYONGGGAFLIPYCYMMLFGFLPFLFELALGQYHNRGCGTLMKRICPALKGVG 122
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 104 RPPYCYONGGGAFLPPTTMAIBGIFLPFMELALGQYHNRGCGTLMKRICPALKGVG 163
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
0Y 123 AICMIDYMGWYNNIIGMAVYYLLIASINSVLPMTSCDNEMWTPLCPTVSPOT--- 179
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 164 AICIAEIAFYASYNTIMAMALYYLL---SPTFDQLPMTSCKNMWTGCTNFTSDDNTW 220
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
0Y 180 NPNSSTPAKEEFERNVLEQHKSNGLDGMGPTRKPSALCAGFVGVYLVYSLMKGVRSACKV 239
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 221 TLHSTPAEEFYTRHVLIDHRSKGLQDGLQISWALCALCMLIFVYFYSIMWGVKTSCKV 280
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
0Y 240 VVVVATALAPYVVLILLARGVTLPGATBESIRYRLPPEMHKLONSKWYIDAAQIFPSLPG 299
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 281 VVVVATAPFYIILSVLVGATLPGAMRQVLYFLKPMQKLTETGVWIDAAQIFPSLPG 340
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
0Y 300 FGTLLASTSYKPNNNCYRDALLITSINCLSPLAGFYFISVLGMVAHVOKSJEVYLE 359
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 341 FGVLLAEFASYKPNNNCYQDALVITSVNCMTSPSGFYIFVVLGMVAHMREDDSEVAKD 400

```

QY 360 -GPGLVFVYPPALATMGSGVMAIFLMLITLGLDSTFGGLEAVTTALCDEYPRVLGR 418
DB 401 AGPSLLFTYPAALANMPASTFEAFIEFLMLITLGLDSTFGGLEAVTTALCDEYPRVLGR 460
QY 419 HREYVAVALLFTYICALPTTYGGVYLVLDLNVGPGALIFVYFAEAGCWYGVDR 478
DB 461 RREYVAVALLFTYICALPTTYGGVYLVLDLNVGPGALIFVYFAEAGCWYGVDR 520
QY 479 FSEDYRTMLGTHPGNFWRTGWSYISPVFLVLFVSVLAHEMLGEGYTPSPMSITVGVN 538
DB 521 FCRDVKEMHGFSPGNFWKICWVAISFLFLITICSLMSPPOURLFQYVYPTWSITLGYC 580
QY 539 MTGTVSCIPLYIYKLLITPGNCINR-1KTQREPTVSIPPADSTL 584
DB 581 ICTSSFCIPYIAYRLITPGTFKRIKISITPPEPTELPQGDURL 627

RESULT 4

US-09-843-598-5
Sequence 5, Application US/09843598
Patent No. US20020010944A1

GENERAL INFORMATION:

APPLICANT: Horvitz, H. Robert
APPLICANT: Ranganathan, Rajesh
TITLE OF INVENTION: CESERT GENES, PROTEINS, AND MODULATORY
TITLE OF INVENTION: COMPOUNDS
FILE REFERENCE: 01997/525002
CURRENT APPLICATION NUMBER: US/09/843,598
CURRENT FILING DATE: 2001-04-26
PRIOR APPLICATION NUMBER: US 60/200,549
PRIOR FILING DATE: 2000-04-26
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 671
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-09-843-598-5

Query Match

Best Local Similarity 50.3%; Score 1579.5; DB 10; Length 671;
Matches 300; Conservative 110; Mismatches 161; Indels 45; Gaps 11;

QY 9 APTAPPDLPATTAKRSRSVVS-----LTPARQ-----RE 39
DB 43 APTA-SEVWPLSADKPLRLVYSTSHSIDPNEPIALGSLTPKBEKRVVALRRSSMVD 101
QY 40 TWAKKAEPFLAVVGAFAVDLGNWREPYICYONGGAFLLPYCYMLLFGPLPFELELAG 99
DB 102 KWATKMEFLAVVGAFAVDLGNWREPYICYONGGAFLLPYCYMLLFGPLPFELELAG 161
QY 100 QYHRCGLTLMKRICPALKGVGATCMIDIMGMNTTIGNAVYLLASLASI-NSVLP 158
DB 162 QYHRCGCVSIVMKVCPFLRGIGYICCTCTFAIFNAIIAQAFAVFAISLSKIMSEVP 221
QY 159 WTSQDNEMWTPCTP---VTSPTQPNNSSTPAKEFEFERNVLEQKNSGLDDMGPIRPSLA 215
DB 222 MASCGPNMTPRCSDDLNTTISRNGRPLTPPSEYLLYVLEQKSTGDDGLGVTSM 281
QY 216 LCVFGEVFLVYSLMKVRSAGKVVWYVATLADYVLLILLARGVTLPGATEGIRYVLP 275
DB 282 VCLLAVFIWYFALMKGPSSGKIYVWATAPYIILSILLRGLLPGAKNGLYYVTPD 341
QY 276 WHKLOSKYVDAASQIFPSLPGFGTLLALSSYNKFNNNCRDALITSINCLTSFLAG 335
DB 342 FEKLDPAAVMSAAQIFPSLPGFGVLLALSSYNFNNNCRDAVTTISINCAVSFSG 401
QY 336 FVIFSVLGYMAHVONKSIE-VGLEBPGLVFIYVPAIATMGVMAIIFFLMLTTGL 394
DB 402 CVFESTLIGMSLLTKNPINEVGEHDASLFIYVPAQALTMQYSCWMSIFEFVMTLIGI 461
QY 395 DSTFGGLEAVTTALCDEYPRVLGRHREYVAVALLFTYICALPTTYGGVYLVLDLNVG 454

DB 462 DSTFAGIEAFITGFCDE-SRFLSKNRKMFVLICIIYFLESPAISYGGQFVIFLDEYX 520
QY 455 PGALIFVYFAEAGCWYGVDRPSEDYRTMLGTHPGNFWKICWVAISFLFLITICSLMSPPOURLFQYVYPTWSITLGYC 514
DB 521 VLSVLFVITCEIAVCMFYGDQFSKDIRAMIGFPGIYWRGWT-CSPVFSVLEIWT 579
QY 515 VLE-AHEEMLGEYTPSPMSITVGVNMTGTVSCIPLYIYKLLITPGNCINR1KTQREPTVSIPPADSTL 572
DB 580 VYNSRKPLOMSYTPRPMWSVILGWFLRLSLVALIPFAIYLLSTGTLYERFRMAITP 639
QY 573 E-----VTSIPPADST 583
DB 640 QQRNRSATSL-AADPT 654

RESULT 5

US-09-843-598-7
Sequence 7, Application US/09843598
Patent No. US20020010944A1

GENERAL INFORMATION:

APPLICANT: Horvitz, H. Robert
APPLICANT: Ranganathan, Rajesh
TITLE OF INVENTION: CESERT GENES, PROTEINS, AND MODULATORY
TITLE OF INVENTION: COMPOUNDS
FILE REFERENCE: 01997/525002
CURRENT APPLICATION NUMBER: US/09/843,598
CURRENT FILING DATE: 2001-04-26
PRIOR APPLICATION NUMBER: US 60/200,549
PRIOR FILING DATE: 2000-04-26
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 671
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-09-843-598-7

Query Match

Best Local Similarity 50.0%; Score 1571.5; DB 10; Length 671;
Matches 299; Conservative 110; Mismatches 162; Indels 45; Gaps 11;

QY 9 APTAPPDLPATTAKRSRSVVS-----LTPARQ-----RE 39
DB 43 APTA-SEVWPLSADKPLRLVYSTSHSIDPNEPIALGSLTPKBEKRVVALRRSSMVD 101
QY 40 TWAKKAEPFLAVVGAFAVDLGNWREPYICYONGGAFLLPYCYMLLFGPLPFELELAG 99
DB 102 KWATKMEFLAVVGAFAVDLGNWREPYICYONGGAFLLPYCYMLLFGPLPFELELAG 161
QY 100 QYHRCGLTLMKRICPALKGVGATCMIDIMGMNTTIGNAVYLLASLASI-NSVLP 158
DB 162 QYHRCGCVSIVMKVCPFLRGIGYICCTCTFAIFNAIIAQAFAVFAISLSKIMSEVP 221
QY 159 WTSQDNEMWTPCTP---VTSPTQPNNSSTPAKEFEFERNVLEQKNSGLDDMGPIRPSLA 215
DB 222 MASCGPNMTPRCSDDLNTTISRNGRPLTPPSEYLLYVLEQKSTGDDGLGVTSM 281
QY 216 LCVFGEVFLVYSLMKVRSAGKVVWYVATLADYVLLILLARGVTLPGATEGIRYVLP 275
DB 282 VCLLAVFIWYFALMKGPSSGKIYVWATAPYIILSILLRGLLPGAKNGLYYVTPD 341
QY 276 WHKLOSKYVDAASQIFPSLPGFGTLLALSSYNKFNNNCRDALITSINCLTSFLAG 335
DB 342 FEKLDPAAVMSAAQIFPSLPGFGVLLALSSYNFNNNCRDAVTTISINCAVSFSG 401
QY 336 FVIFSVLGYMAHVONKSIE-VGLEBPGLVFIYVPAIATMGVMAIIFFLMLTTGL 394
DB 402 CVFESTLIGMSLLTKNPINEVGEHDASLFIYVPAQALTMQYSCWMSIFEFVMTLIGI 461
QY 395 DSTFGGLEAVTTALCDEYPRVLGRHREYVAVALLFTYICALPTTYGGVYLVLDLNVG 454
DB 462 DSTFAGIEAFITGFCDE-SRFLSKNRKMFVLICIIYFLESPAISYGGQFVIFLDEYX 520


```

Db      365 KTFVSGSESEYFYFLK--ISAGIEYPEGRMPLALCLFLAWVIYASLAKIKTSKVV 422
Qy      241 WYATALAPYVLLILLARCVTLPGATEGIKYVLTPEMHKQNSKWIDAASQIFPSLGP 300
Db      423 YFTATFPYVLLILLIRGVTLPAGAGIWFYFTPKWEKLTNATWKDATQIFPSLSAM 482
Qy      301 GTLLALSSYNKNNCYRDALITSSINCLTSFLAGVIFPSVLGYMAHVONKISIEVGL 360
Db      483 GGLILSSYNKNNHCYRDLTCTNSATSIAGVIFPSVLGYMAHVONKISIEVGL 542
Qy      361 POLVIVPEALVATMTGVSFAIIFFLMILTLGDSFEGLEAVTTALCDEXPRVLR 420
Db      543 POLVIVPEALVATMTGVSFAIIFFLMILTLGDSFEGLEAVTTALCDEXPRVLR 602
Qy      421 EYFVAVLLFIYICALPTTGGVYVLDLNVYGCALFLFVPAEAGVGVGVDRS 480
Db      603 PFTTGCCCFETIMFPMATGGITMFOVLVDYAAVALYIAIFELVGVISTVYGLQRC 662
Qy      481 EDVRLMHTPGMFWRTGMSYISPVFLVLFVSVLAHEMLGCEYTPSMITVGMWT 540
Db      663 EDIEMWIGQPNIFMKVCAEFTPLIFLFCFSFYQWMPMTYGSVRYPMNSMVLGML 722
Qy      541 GTTVCSCIPLYIYIKLLITPGNCINRIKTQRP 573
Db      723 ACSVIMIPIMFYIKMHLAPGRFIERKLKACSPQ 755

RESULT 10
US-09-815-923-14
; Sequence 14, Application US/09815923
; Publication No. US20020197644A1
; GENERAL INFORMATION:
; APPLICANT: Gill, Sarjeet S.
; APPLICANT: Ross, Linda S.
; TITLE OF INVENTION: The Regents of the University of California
; TITLE OF INVENTION: Use of Insect Cell Membrane Transporters as No. US20020197644A1
; FILE REFERENCE: 023070-093800US
; CURRENT APPLICATION NUMBER: US/09/815,923
; CURRENT FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 14
; LENGTH: 597
; TYPE: PRT
; ORGANISM: Manduca sexta
; FEATURE:
; OTHER INFORMATION: GABA transporter
; 09-815-923-14

Query Match      41.5%; Score 1302; DB 11; Length 597;
Best Local Similarity 42.5%; Pred. No. 1.6e-108;
Matches 248; Conservative 108; Mismatches 183; Indels 44; Gaps 8;

Qy      18 PATTAOKRSRVVSLTPARQRETAKKAFFLLAVGVAVDLGNVWRPFYICQNGGAF 77
Db      23 PSDVAVKSN-----LPERGNSAKSLDFILSVIGLIGGNVWRPFYICKNGGAF 74
Qy      78 IPYCVMLLFGSLPFLFELALAGYHRCCLLMKRICALGVGAYATIMIDYMGMYNT 137
Db      75 IPYCVMLLFGSLPFLFELALAGYHRCCLLMKRICALGVGAYATIMIDYMGMYNT 133
Qy      138 IIGNAVYLLASLINSVLPMTSCDNEMNTPLCTPVTSPQNTNP----- 182
Db      134 ILAAHIFEFMS--KSDVPMRNDNTWNTATC--VNPYDRKMLTJWCSSLDGNSFTCT 187
Qy      183 -----STPAKEFEERNVLEOHKNSGLDDMGPIKPSLALCVGFVLYVSLMKG 232
Db      188 LNRNVSAKAVSDPYKKEWERRALQ--ISSGIEHGNIRNMLAGLLLVWVLCYCIK 245
Qy      233 VRSAGKYVWATALAPYVLLILLARCVTLPGATEGIRYVLTPEMHKQNSKWIDAASQ 292

```

```

Db      246 VMTGKVYVYFALPFYLLTVLLIRGILTLPAGAMEGIKFVMDPMSKLESEVVIDAVTQ 305
Qy      293 FESLGPGETLLALSSYNKNNCYRDALITSSINCLTSFLAGVIFPSVLGYMAHVONK 352
Db      306 FESYGLGDTLLALSSYNKNNCYRDALITSSINCLTSFLAGVIFPSVLGYMAHVONK 365
Qy      353 IEEVGLGGLVFIYPEALVATMTGVSFAIIFFLMILTLGDSFEGLEAVTTALCDEX 412
Db      366 VAEVASAGGLAFVAVSADLPAPALMSCLFFMILLIGDSQFCMTEGITVAIDEM 425
Qy      413 PVILGHRFVAVALLFIYICALPTTGGVYVLDLNVYGCALFLFVPAEAGVGVDRS 471
Db      426 PKLRRRKEIFAIICIIISLVGLCISEGMYVQIIDSVAVSGCLLIFFECSVIS 485
Qy      472 WYGVDRSESDVRLMHTPGMFWRTGMSYISPVFLVLFVSVLAHEMLGCEYTPSM 531
Db      486 WAFGVNRFYDGIKEMIGITPTIMMFCWVGFTPAICISIFITNLQWMPRTKMYNEYPM 545
Qy      532 SITVGMVMTGTVSCIPLYIYIKLLITPGNCINRI-KTQORPE 573
Db      546 SHAFGMFTALSSMLCIPGYMIYLMVTRGTQWKEFKHIVRIPE 588

RESULT 11
US-09-919-039-378
; Sequence 378, Application US/09919039
; Publication No. US20030108871A1
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 378
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No. US20030108871A1 5834958CD1
; US-09-919-039-378

Query Match      39.7%; Score 1247.5; DB 12; Length 614;
Best Local Similarity 41.1%; Pred. No. 1.3e-103;
Matches 242; Conservative 116; Mismatches 208; Indels 23; Gaps 9;

Qy      7 PPAPTAPPDDLPTTAOKRSRVVSLTPARQRETAKKAFFLLAVGVAVDLGNVWRPFY 66
Db      12 PPAVSWVEBECKLDDQEDDOV-----KDRGQNTKMEFVLSVAGELIGLVNWRPFY 64
Qy      67 ICYONGGAFLLIPYCVMLLFGSLPFLFELALAGYHRCCLLMKRICALGVGAYATIM 126
Db      65 LCKYNGGAFLLIPYCVMLLFGSLPFLFELALAGYHRCCLLMKRICALGVGAYATIM 124
Qy      127 IDIYGMVYNTIIGNAVYLLASLINSVLPMTSCDNEMNTPLCTPVTSPQ---TNP- 181
Db      125 IESYLVNYYIIILAAWLFY--FSSFTSELPTTCNNFMWTEHCTDCLNHSAGATVP 181
Qy      182 -NSSTPAKEFEERNVLEOHKNSGLDDMGPIKPSLALCVGFVLYVSLMKGVSAGVY 240
Db      182 EFTSPVKEFERRVL--GITSIHDLGSLMELALCLLLAMVICYFCITMGVASTGVY 239
Qy      241 WYATALAPYVLLILLARCVTLPGATEGIRYVLTPEMHKQNSKWIDAASQIFPSLGP 300
Db      240 YFTATFPYVLLILLIRGVTLPAGAGIWFYFTPKWEKLTNATWKDATQIFPSLSAM 299
Qy      301 GTLLALSSYNKNNCYRDALITSSINCLTSFLAGVIFPSVLGYMAHVONKISIEVGL 360
Db      300 GCLTALAGSYNKNYHNCYKDCIALCLFNSATSEFVAGFVFSITGPMDSQGVPISEVAESG 359

```


